CMS Physics Analysis Summary

Contact: cms-pag-conveners-higgs@cern.ch 2021/09/18

A search for light Higgs bosons from supersymmetric cascade decays in proton-proton collisions at 13 TeV

The CMS Collaboration

Abstract

A search for pairs of boosted light Higgs bosons (H_1) produced in supersymmetric cascade decays is performed in final states with small missing transverse momentum. The complete LHC Run II proton-proton collision data set is used, recorded with the CMS detector at a centre-of-mass energy of 13 TeV and corresponding to an integrated luminosity of 137 fb^{-1}. The search targets events where both H_1 bosons decay into $b\bar{b}$ pairs that are reconstructed as large-radius jets using substructure techniques. No evidence is found for any excess of events beyond the background expectations of the standard model (SM). Results from the search are interpreted in the next-to-minimal supersymmetric extension of the SM, where a low-mass singlino leads to multi-step squark and gluino decays that can predominantly end with a boosted singlet-like H_1 boson and a low-momentum singlino-like neutralino. Upper limits are set on the product of the squark or gluino pair-production cross section and the $b\bar{b}$ branching ratio of the H_1 for a benchmark model with almost mass-degenerate light flavour squarks and gluinos. Under the assumption of an SM-like $H_1 \to b\bar{b}$ branching ratio, H_1 bosons with masses in the range 40–120 GeV, arising from the decays of squarks or gluinos with a mass from 1200–2500 GeV, are excluded at the 95% confidence level.

1 Introduction

This note presents a search for pairs of boosted light Higgs bosons (H_1) produced in supersymmetric (SUSY) [1–8] cascade decays in final states with small missing transverse momentum (p_T^{miss}). These events can be the primary signature of pair-production of squarks ($\tilde{q}$) and gluinos ($\tilde{g}$) in the next-to-minimal supersymmetric extension of the standard model (NMSSM) [9] when the lightest SUSY particle (LSP) is a low-mass singlino-like neutralino ($\tilde{\chi}_S^0$) [10]. A $\tilde{\chi}_S^0$ LSP has only small couplings to the other SUSY particles, suppressing direct squark or gluino decays to the $\tilde{\chi}_S^0$ and leading to multi-step squark and gluino cascade decays ending with the $\tilde{\chi}_S^0$ and a Higgs, Z, or W boson [10, 11]. We focus on the case of a CP-even H_1 boson that is mostly an SM gauge singlet (Figure 1). When the $\tilde{\chi}_S^0$ has a much smaller mass than the H_1 and the phase space for the next-to-LSP (NLSP) $\tilde{\chi}_2^0 \to H_1 + \tilde{\chi}_S^0$ decay is small, the H_1 carries much larger momentum than the $\tilde{\chi}_S^0$ [10]. In this p_T^{miss}-supressed scenario, the key signature of the pair-production of squarks and gluinos is a pair of boosted H_1 bosons.

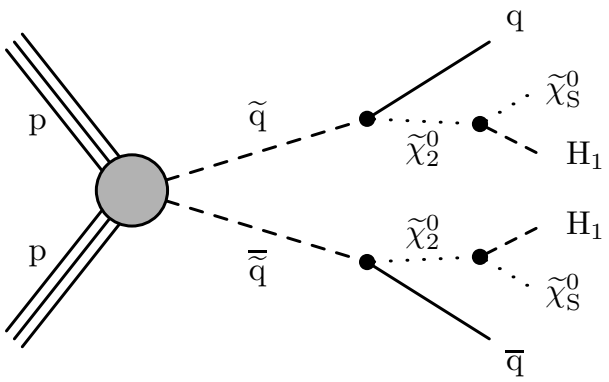

Figure 1: Diagram of squark pair production and subsequent cascade decay in the benchmark signal model. The particle $\tilde{\chi}_2^0$ is the NLSP, $\tilde{\chi}_S^0$ is the singlino-like LSP, and H_1 is the CP-even singlet-like Higgs boson.

This search targets events with two boosted H_1 bosons that decay into $b\bar{b}$ pairs that are reconstructed as large-radius jets using substructure techniques. It is the first LHC search to focus on such events with small p_T^{miss}, which evade constraints from p_T^{miss}-based searches [10, 12]. The analysis uses a data sample of proton-proton (pp) collision events collected by the CMS experiment at centre-of-mass energy $\sqrt{s} = 13\,\text{TeV}$ in the years 2016–2018, corresponding to an integrated luminosity of $137\,\text{fb}^{-1}$ [13–15].

2 Benchmark signal model

We establish a benchmark signal model following those of Refs. [10, 11]. The eight light-flavour squarks are assumed mass-degenerate, with mass M_{SUSY}, and the gluino mass is set 1% larger than that of the squarks. In the $\tilde{q}\tilde{g}$ and $\tilde{g}\tilde{g}$ production modes, the gluinos decay as $\tilde{g} \to \tilde{q} + q$. The squarks then decay as they do in Figure 1. The nominal gluino-squark mass gap means that, for a given squark mass, the total production cross section is almost maximal whilst still allowing the $\tilde{g} \to \tilde{q} + q$ decay to occur. Because the mass gap is so small, only a small amount of momentum is transferred to the quark in the gluino decay, and the kinematics of the final-state particles are very similar in the $\tilde{q}\tilde{q}$, $\tilde{q}\tilde{g}$, and $\tilde{g}\tilde{g}$ production modes.

The search targets squarks and gluinos with masses $M_{SUSY} \geq 1200\,\text{GeV}$. Less massive squarks

and gluinos would result in final states with lower momentum H_1 bosons decaying into $b\bar{b}$ pairs with larger angular separation that could exceed the radius used for jet reconstruction (see Section 4). The signal cross sections for the M_{SUSY} values probed in this search, calculated at approximately NNLO+NNLL (next-to-next-to-leading order and next-to-next-to-leading logarithm) in α_S [16–24], are shown in Table 1.

Table 1: The M_{SUSY} values considered in this search and corresponding production cross sections (the sum of $\tilde{q}\tilde{q}$, $\tilde{q}\tilde{g}$, and $\tilde{g}\tilde{g}$), calculated at approximately NNLO+NNLL in α_S [16–24] for a squark mass of M_{SUSY} and a gluino mass 1% larger. The quoted uncertainty is from variations in the choice of scale, parton distribution functions, and α_S.

M_{SUSY}	$\sigma(pp \to \tilde{q}\tilde{q}, \tilde{q}\tilde{g}, \tilde{g}\tilde{g})$ [pb]	Uncertainty
1200 GeV	0.57940	7.9%
1600 GeV	0.06882	9.4%
2000 GeV	0.01013	11.4%
2200 GeV	0.00405	12.7%
2400 GeV	0.00164	14.3%
2600 GeV	0.00067	16.0%
2800 GeV	0.00027	18.2%

The different values of the H_1 mass (M_{H_1}) considered, along with the corresponding branching ratios of $H_1 \to b\bar{b}$, are shown in Table 2. Only events where both H_1 bosons decay into $b\bar{b}$ pairs are considered as signal. The H_1 branching ratios are chosen to be close to those of a SM-like Higgs boson of the corresponding mass [10], and were calculated using the HDECAY package [25]. The branching ratio to $b\bar{b}$ decreases for larger H_1 masses as the WW^* and ZZ^* decay channels, both of which have sizable leptonic branching fractions, become more accessible. The region $M_{H_1} < M_Z$ is therefore where the p_T^{miss}-suppressed all-hadronic signature is of most experimental interest. Nevertheless, to preserve generality in this analysis, the search attempts to probe as much of the region $M_{H_1} \leq 125\,GeV$ as possible.

Table 2: The M_{H_1} values considered in this search and corresponding $H_1 \to b\bar{b}$ branching ratios, calculated for a SM-like Higgs boson using the HDECAY package [25].

M_{H_1}	BR ($H_1 \to b\bar{b}$)
30 GeV	0.868
35 GeV	0.867
40 GeV	0.865
50 GeV	0.858
60 GeV	0.850
70 GeV	0.840
80 GeV	0.829
90 GeV	0.816
100 GeV	0.795
110 GeV	0.749
125 GeV	0.581

In addition to M_{H_1} and M_{SUSY}, there are two other unknown masses in the signal model: those of the $\tilde{\chi}_S^0$ and the $\tilde{\chi}_2^0$. The corresponding degrees of freedom are parameterised by $R_M \equiv M_{H_1}/M_{\tilde{\chi}_2^0}$ and $\Delta_M \equiv M_{\tilde{\chi}_2^0} - M_{H_1} - M_{\tilde{\chi}_S^0}$. The p_T^{miss}-suppressed signature arises for values of R_M close to unity, provided Δ_M is sufficiently small that the $\tilde{\chi}_2^0 \to H_1 + \tilde{\chi}_S^0$ decay is possible. In this scenario the phase space for the $\tilde{\chi}_2^0$ decay is small and the $\tilde{\chi}_S^0$ has much smaller

mass than the H_1, such that the $\widetilde{\chi}^0_S$ always carries much less momentum than the H_1. For the benchmark model, we assume $R_M = 0.99$ and $\Delta_M = 0.1\,\text{GeV}$. For the values of M_{H_1} and M_{SUSY} considered, the visible final-state kinematics are not strongly dependent on these choices in the region $R_M > 0.9$. The p_T^{miss}-suppressed signature probed in this search is therefore representative of a significant part of the model parameter space. For smaller R_M, p_T^{miss}-based searches would become sensitive [10, 12].

We assume that no particles other than the gluinos and those shown in Figure 1 play any role in the production or decay, and that the decays $\widetilde{q} \to q + \widetilde{\chi}^0_2$ and $\widetilde{\chi}^0_2 \to H_1 + \widetilde{\chi}^0_S$ occur with branching ratios of unity. This is true in the NMSSM in the R_M and Δ_M region of interest, except where $M_{\widetilde{\chi}^0_2} > M_Z + M_{\widetilde{\chi}^0_S}$. In that case, the $\widetilde{\chi}^0_2 \to Z + \widetilde{\chi}^0_S$ decay is allowed if the $\widetilde{\chi}^0_2$ has a higgsino component [11]. However, the $\widetilde{\chi}^0_2$ is expected to be mostly bino-like for relevant values of its mass [10]. For configurations where the H_1 mass is close to that of the SM-like Higgs boson (H_{SM}), the decay $\widetilde{\chi}^0_2 \to H_{SM} + \widetilde{\chi}^0_S$ is also possible. The signatures of such H_1 and H_{SM} bosons would be indistinguishable in this search, and the results can be interpreted under the assumption that the corresponding $\widetilde{\chi}^0_2$ branching fractions sum to unity.

3 Event simulation

The primary background in this search is from multijet production. Simulated multijet events are used to validate the multijet background estimation from data (described in Section 5), but are not used for any of the final predictions. The yields from $t\bar{t}$, Z+jets and W+jets production, which represent a background when the vector bosons decay into quark anti-quark pairs, are determined from simulated events.

The multijet, Z+jets and W+jets processes are simulated at leading order (LO) in perturbative quantum chromodynamics (QCD) using MADGRAPH5_aMC@NLO 2.4.2 [26] with up to four additional partons at the matrix element (ME) level. Simulated signal events for each pair of M_{H_1} and M_{SUSY} values of the benchmark model, introduced in Section 2, are generated at LO at the ME level with up to one additional parton using MADGRAPH5_aMC@NLO 2.3.3. The MLM [27] prescription is used to match partons from the LO matrix element calculations to those from the parton showers. Simulated $t\bar{t}$ events are produced at next-to-leading order (NLO) in perturbative QCD at the ME level by the POWHEG v2.0 [28–31] generator. The NNPDF2.3, NNPDF3.0, and NNPDF3.1 [32–35] parton distribution function (PDF) sets are used for the signal, 2016 background, and 2017–2018 background simulations, respectively, and the parton shower and hadronisation are performed by PYTHIA 8.2 [36]. The CUETP8M1 [37, 38] tune is used for the signal and 2016 background simulations, while the CP5 tune [39] is used for the 2017 and 2018 background simulations. The cross sections used to normalise the $t\bar{t}$, Z+jets and W+jets simulations are calculated at NNLO in perturbative QCD [40–43]. Additional pp interactions within the same or nearby bunch crossings (pileup) are simulated for all events according to the observed distribution of the number of interactions in each bunch crossing [44]. The interactions of particles with the CMS detector are simulated using GEANT4 [45].

4 Event Selection

The data are collected using triggers [46, 47] based on the scalar sum of jet transverse momentum (H_T), with a requirement of $H_T > 900\,\text{GeV}$ (2016) and $H_T > 1050\,\text{GeV}$ (2017–18). Since the final state is hadronic only, the typical signal event H_T depends significantly on M_{SUSY}, and signal events with $M_{SUSY} \geq 1200\,\text{GeV}$ tend to exhibit $H_T > 1500\,\text{GeV}$ (see Figure 2). A minimum

H_T of 1500 GeV is required in the kinematic event selection defined below. Although the offline H_T resolution is better than that at trigger level, this is sufficiently beyond the trigger-level requirement to result in a trigger efficiency of 100%, as measured using events collected with an approximately independent muon-based trigger.

The events selected by the trigger system are reconstructed offline using a particle-flow algorithm [48], which aims to reconstruct and identify each individual particle in an event using an optimised combination of information from the elements of the CMS detector.

Jets are reconstructed by clustering the particle-flow candidates using the anti-k_T clustering algorithm [49, 50]. Clustering is done with a distance parameter of 0.8 (0.4) for large (standard)-radius jets, referred to as AK8 (AK4) jets. The jet momentum is determined as the vectorial sum of all particle momenta in the jet. The pileup per particle identification algorithm (PUPPI) [51, 52] is used to mitigate the effect of pileup at the reconstructed particle level, making use of local shape information, event pileup properties, and tracking information. Jet energy corrections are derived from simulation to bring the measured response of jets to that of particle level jets on average. In situ measurements of the momentum balance in dijet, photon + jet, Z + jet, and multijet events are used to account for any residual differences in the jet energy scale and resolution between data and simulation [53, 54]. Additional selection criteria are applied to each jet to remove jets potentially dominated by instrumental effects or reconstruction failures.

The identification of AK8 jets originating from two collimated b quarks (double-b tagging) is integral to the reconstruction of the H_1 bosons. A discriminator value is calculated for each jet using a double-b tagging algorithm that combines tracking and vertexing information in a multivariate approach with no significant dependence on jet mass or p_T [55]. The efficiency for selecting jets originating from boosted $H_1 \to b\bar{b}$ decays is about 75%, while the efficiency for jets containing only light flavour quarks or gluons is about 5%.

The event pre-selection requires two AK8 jets with transverse momentum $p_T > 170$ GeV and $|\eta| < 2.4$ (so that they are within the acceptance of the tracker). If there are more than two candidate AK8 jets, the two with the highest double-b-tag discriminator values are selected as those most likely originating from $H_1 \to b\bar{b}$ decays. The two selected AK8 jets are then randomly assigned the labels 'A' and 'B'. The H_T distributions for various simulated signal and background processes are shown in Figure 2, after application of the pre-selection.

Additional requirements based on the expected kinematic properties of signal events are applied after pre-selection. They define the kinematic event selection:

- Both of the selected AK8 jets must have $p_T > 300$ GeV and $|\eta| < 2.4$, typical of the jets originating from $H_1 \to b\bar{b}$ decay.
- There must be at least one AK4 jet with $p_T > 300$ GeV and $|\eta| < 3.0$, typical of the quarks produced in the squark decays illustrated in Figure 1. It must be separated by $\Delta R \equiv \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2} > 1.4$ from both of the selected AK8 jets, to avoid being composed of the same particle-flow candidates.
- The event H_T must exceed 1500 GeV, characteristic of the high-momentum all-hadronic final state in signal events (see Figure 2). For the offline analysis, we define $H_T \equiv \sum_{\text{AK4 jets}} p_T$, using all AK4 jets with $p_T > 40$ GeV and $|\eta| < 3.0$ (including AK4 jets with particle-flow candidates also clustered into AK8 jets).

For the 2018 data set only, the $|\eta|$ selection for the AK4 jets is reduced from 3.0 to 2.4 to avoid jets whose energy is affected by noise resulting from the crystal transparency losses in the endcap electromagnetic calorimeters at large $|\eta|$ [56]. This change has a negligible effect on signal

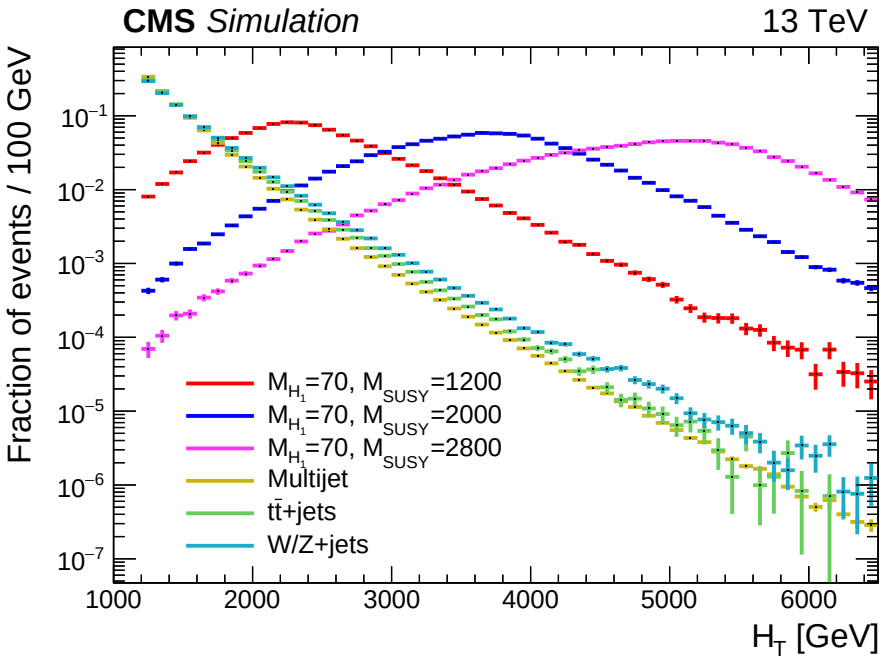

Figure 2: The normalised H_T distribution in signal events with different values of M_{SUSY}, and for the simulated SM background processes, as labelled in the legend. The numbers specified in the legend have unit GeV. All events satisfy the pre-selection and have $H_T \geq 1200$ GeV.

efficiency for all mass hypotheses considered. The fraction of signal events that satisfy the kinematic event selection is approximately independent of M_{H_1}. It increases from about 60% to 80% as M_{SUSY} increases from 1200–2000 GeV, after which it remains approximately constant.

4.1 Double-b-tag based event selection

After kinematic event selection, events contain two AK8 jets that have been classified as the boosted $H_1 \to b\bar{b}$ candidates. The double-b-tag discriminator values of the two randomly allocated AK8 jets ('A' and 'B') define a 2D parameter space, as shown in Figure 3 for simulated signal and multijet events. The signal events are expected to contain two $H_1 \to b\bar{b}$ decays and therefore cumulate in the region where both double-b-tag discriminator values are large. The tag region (TR) is defined as the region where the sum of the two double-b-tag discriminator values exceeds 1.3, illustrated by the green shaded triangle in Figure 3. Two additional regions are defined in Figure 3: the control region (CR), a multijet-dominated region with negligible signal contamination; and the validation region (VR), a more signal-like region where one of the two jets has a large double-b-tag discriminator value, but defined sufficiently far from the TR that the signal contamination remains negligible.

After the kinematic event selection has been applied, about 50% of the remaining signal events populate the TR, with variation at the level of $\pm 10\%$ across the considered M_{H_1} and M_{SUSY} parameter space. Since the multijet background is dominated by light flavour quark and gluon initiated jets, only about 3% of the events that satisfy the kinematic selection populate the TR.

4.2 Soft-drop mass based signal and sideband regions

In signal events, both selected AK8 jets are likely to originate from $H_1 \to b\bar{b}$ decays and thus have a jet mass close to M_{H_1}. In the multijet background there is no resonant mass peak, while the other backgrounds are only expected to exhibit peaks near the known top quark and vector

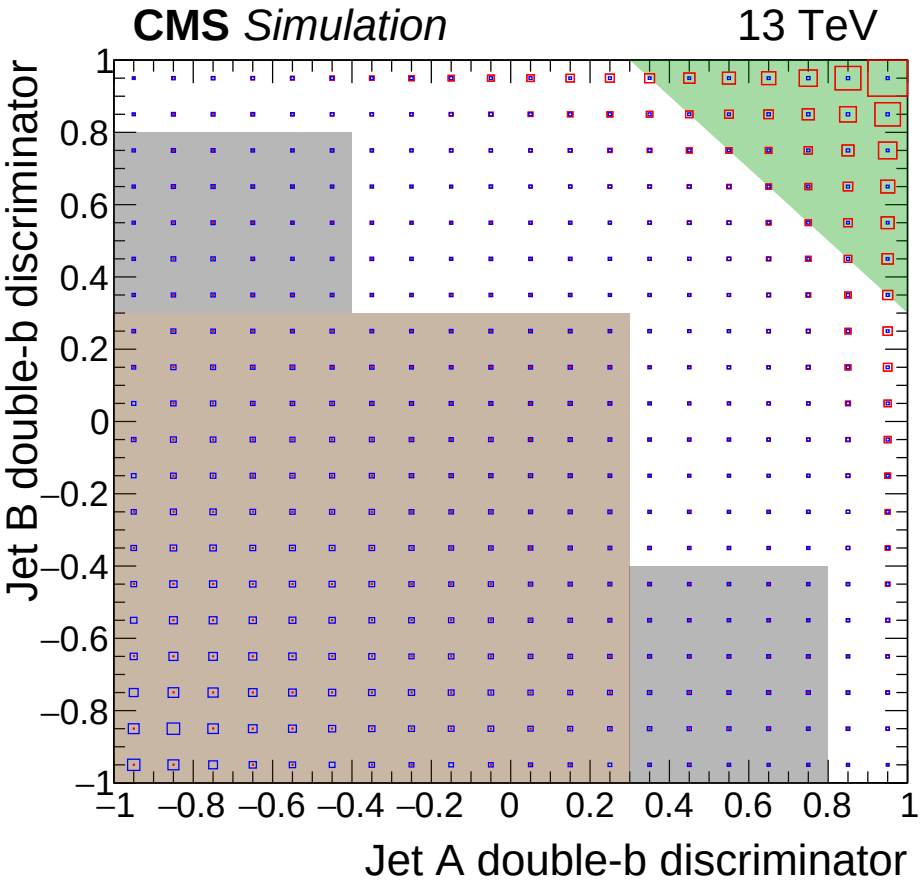

Figure 3: The normalised distributions of simulated signal and multijet events in the 2D double-b-tag discriminator plane, where the event densities in each bin are represented by the areas of red and blue squares, respectively. The signal parameters are $M_{\mathrm{H}_1} = 70$ GeV and $M_{\mathrm{SUSY}} = 2000$ GeV. The kinematic event selection is applied, and the masses of the two AK8 jets are required to be within the set of signal and sideband mass regions defined in Figure 4. The green, brown, and grey shaded areas represent the tag region (TR), control region (CR), and validation region (VR), respectively.

bosons masses, so an accurate reconstruction of the jet mass is vital in distinguishing signal from background. The AK8 jet masses are evaluated using the "soft drop" algorithm [57, 58] (with soft drop threshold $z_{\text{cut}} = 0.1$ and angular exponent $\beta = 0$), in which wide-angle soft radiation is recursively removed from a jet.

The soft-drop masses of the two AK8 jets define a 2D parameter space, shown in Figure 4, in which 10 signal regions S_i and 10 sideband regions U_i are defined. The S_i regions contain events in which the two H_1-candidate jets have approximately the same soft-drop mass. The event distributions for a set of signal models with different M_{H_1} values are shown in Figure 5, with the signal and sideband mass regions overlaid. The widths of the S_i regions are driven by the experimental soft-drop mass resolution at the different simulated values of M_{H_1}. The peaks in the signal distributions where one or both AK8 jets have a soft-drop mass close to zero are a result of either a jet originating from a single parton being mistakenly selected (as one of the two jets with largest double-b-tag discriminator), or one of the $H_1 \to b\bar{b}$ decays lying outside the acceptance of the jet reconstruction algorithm. This can happen when the angular separation ΔR of the b quarks exceeds the AK8 jet distance parameter, or as a result of the z_{cut} threshold in the soft drop algorithm. For signal models with $40 \leq M_{H_1} \leq 125\,\text{GeV}$, approximately 50% of the events that satisfy the kinematic and TR selection fall within the collection of S_i mass regions. However, for $M_{H_1} \leq 35\,\text{GeV}$ the bulk of the distribution falls below the lowest signal mass region, leading to a rapid decrease in signal acceptance.

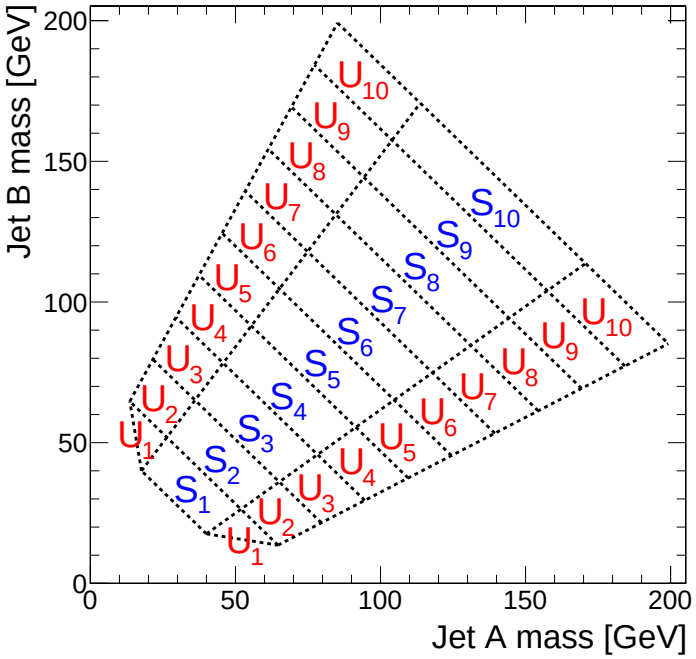

Figure 4: The mass regions used in the 2D soft-drop mass plane. The regions labelled S_i are the signal mass regions, and the disconnected regions U_i form the corresponding sidebands.

The distribution of the simulated multijet events, the dominant background process, is also shown in Figure 5. For the majority of events at least one of the AK8 jets is evaluated to have a small soft-drop mass, since quark and gluon jets are characterised by a one-prong structure. After applying the kinematic and TR event selection criteria, approximately 5% of the remaining multijet events fall within the set of S_i mass regions, with a greater density of events in the lower mass regions.

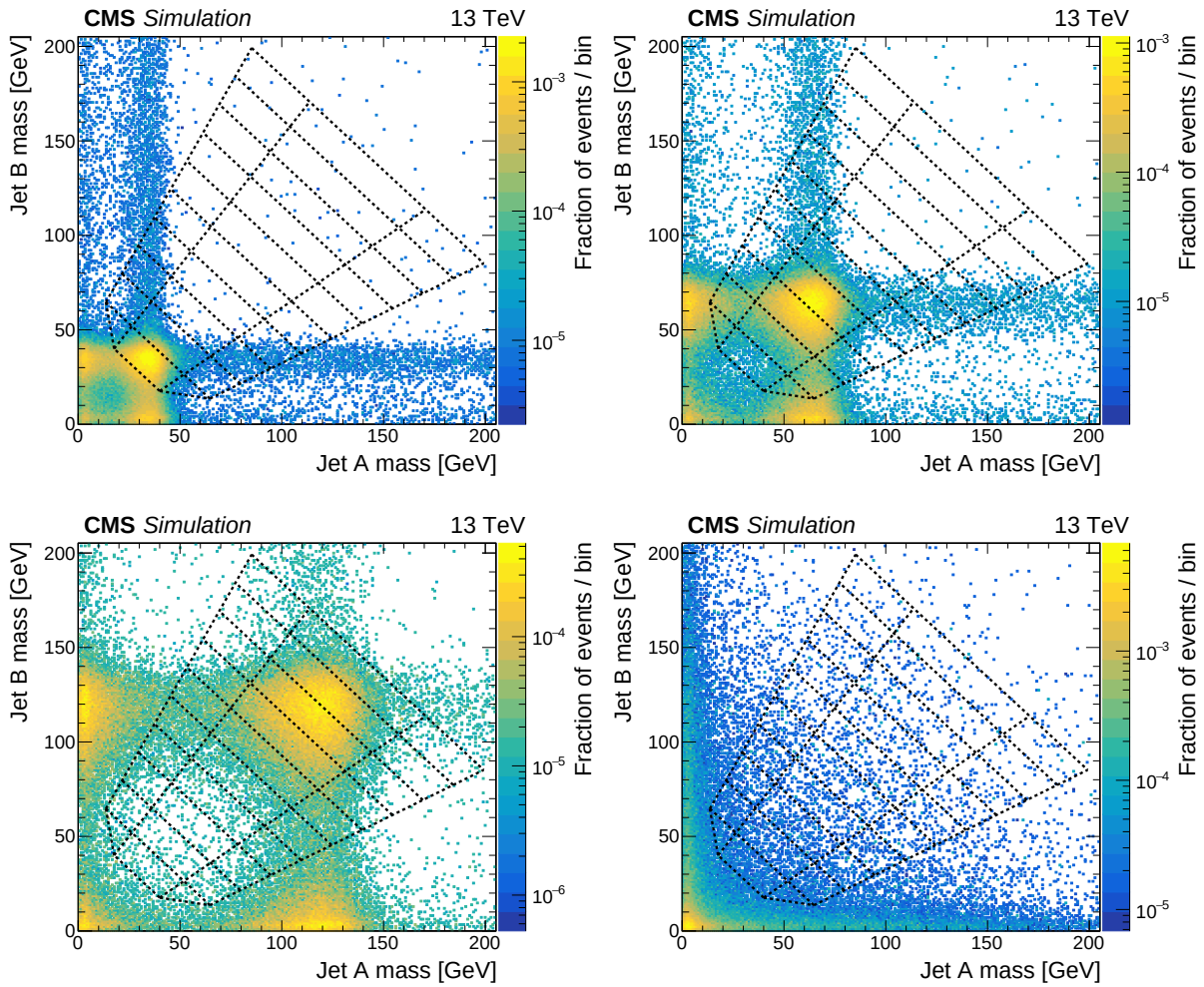

Figure 5: The normalised distribution of events in the 2D soft-drop mass plane with the mass regions overlaid. The first three sub-figures correspond to signal events with $M_{\mathrm{SUSY}} = 2000$ GeV and M_{H_1} values of 40, 70, and 125 GeV respectively. The final sub-figure corresponds to simulated multijet events. All events satisfy the TR requirement and the kinematic event selection.

For each S_i region there is a corresponding sideband region, U_i, used for the data driven multijet background estimation discussed in Section 5. For $i > 1$, the area within the sideband region U_i is the same as that in the corresponding signal region S_i. The sideband regions U_1 take on a triangular shape to avoid the region of very small soft-drop masses (where the density of multijet events increases sharply).

4.3 H_T binning and predicted search region yields

After application of the kinematic and double-b-tag event selection, the events are classified in three H_T regions: 1500–2500, 2500–3500, and 3500+ GeV. Each H_T bin is divided into the 10 mass signal regions S_i defined in Figure 4, resulting in a total of 30 search regions for each data-taking year. The search region yields are visualised in a 30-bin histogram, where bins 1–10 represent the S_i regions, in ascending order, for the first H_T bin ($1500 \leq H_T < 2500$ GeV). The subsequent two sets of 10 bins represent the same for the second ($2500 \leq H_T < 3500$ GeV) and third ($H_T \geq 3500$ GeV) H_T bins.

This 1D representation is used in Figure 6 to compare the observed yields to those of the predicted backgrounds and various signal models. The primary background is from multijet events, estimated from data using the method described in Section 5. The predicted contribution from the $t\bar{t}$ process is also significant, particularly in the larger soft-drop mass regions populated by jets composed of hadronic decay products of top quarks or W bosons. The yields from the Z+jets and W+jets processes are small in comparison. All expected SM background processes tend to exhibit low values of H_T compared to the signal.

The distributions of signal events with $M_{H_1} = 70$ GeV and $M_{SUSY} = 1200$, 2000 and 2800 GeV are also shown in Figure 6. Signal events with larger M_{SUSY} preferentially populate the higher H_T bins, although the production cross section decreases quickly with increasing M_{SUSY}. Within each H_T bin, the distribution of events in the different mass regions is described by a peak with a width of about three bins, centred near the model value of M_{H_1}.

5 Multijet background estimation from data

The mass sideband regions U_i form the basis of the data-driven multijet background yield estimation. Empirically, the multijet event density is found to be approximately uniform within each of the 10 mass regions illustrated in Figure 4 (spanning S_i and U_i for each region i). The sideband regions are constructed to have the same area as the corresponding signal regions, so that the signal region multijet event yields $\hat{S}_i$ are approximately equal to the sideband yields $\hat{U}_i$. The ratios of signal and sideband region yields F_i are measured in the CR (defined in Figure 3) and found to be close to unity except for $F_1 \approx 1.5$, which is larger because the area of region U_1 is smaller than that of S_1.

The multijet background yield in the TR is estimated independently for each signal region S_i using Equation 1:

$$\hat{S}_i^{TR} = F_i \cdot \hat{U}_i^{TR}, \tag{1}$$

where $\hat{U}_i^{TR}$ is the observed TR yield in sideband region U_i after subtracting the contributions from the other simulated backgrounds (Section 3), with the constraint $\hat{U}_i^{TR} > 0$.

Since the factors F_i are measured and applied in different regions of double-b-tag discriminator parameter space, any correlation between the soft-drop mass and the double-b-tag discriminator of the AK8 jets could bias the prediction of Equation 1. Using data, this correlation was investigated and found to have less than a 10% impact on the estimated yields.

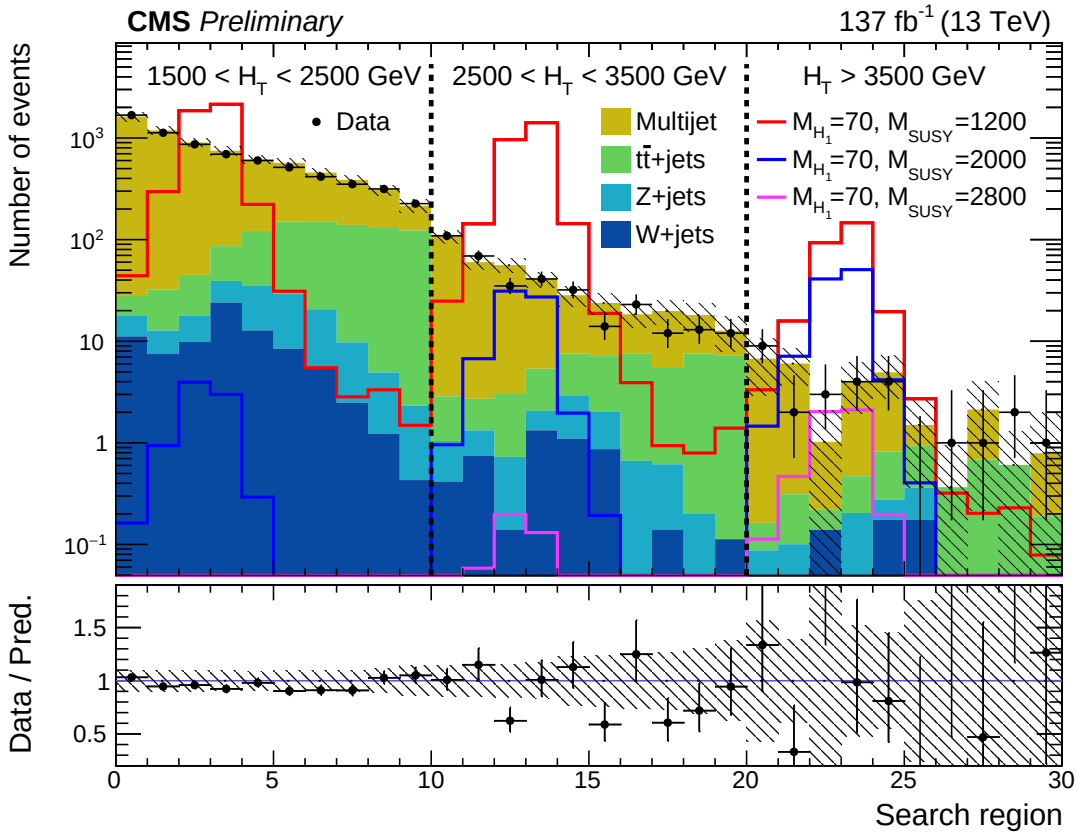

Figure 6: Observed and predicted yields in the search regions, summed over the three data-taking years. The multijet background is from the data-driven prediction of Section 5, while the other backgrounds are from simulation. Example signal distributions are shown for $M_{\mathrm{H}_1} = 70\,\mathrm{GeV}$ and $M_{\mathrm{SUSY}} = 1200$, 2000 and 2800 GeV. The numbers specified in the legend have unit GeV. The error bars represent the statistical uncertainties, and the hatched bands the systematic uncertainties.

The overall accuracy of the multijet estimation method is assessed through closure tests. First, the method was applied to simulated multijet events in the TR. Within the statistical uncertainties, there was no evidence of bias and the predicted yields matched the observed yields for each data-taking year. Second, the method was applied in the multijet-dominated VR in data (defined in Figure 3) by making the appropriate modification to Equation 1: $\hat{S}_i^{\mathrm{VR}} = F_i \cdot \hat{U}_i^{\mathrm{VR}}$. The resulting predicted and observed VR yields are consistent within uncertainties, as shown in Figure 7.

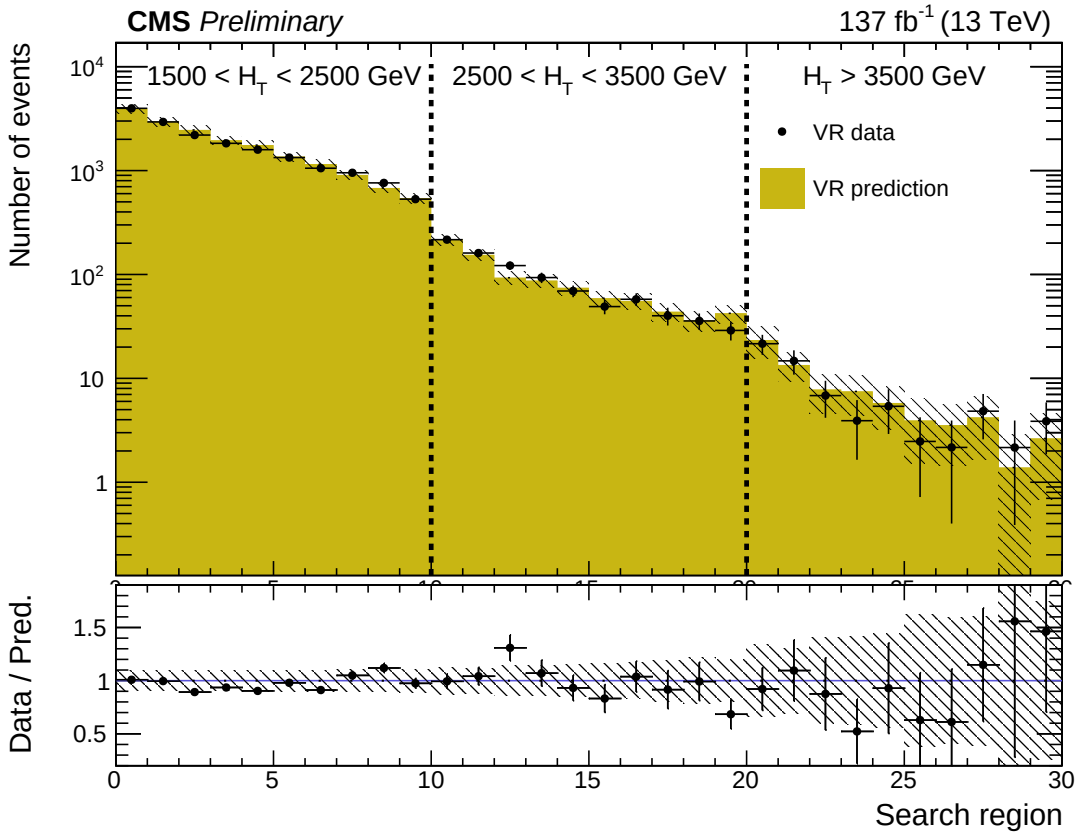

Figure 7: A comparison of the predicted and observed multijet yields in the validation region (VR), after subtraction of the other simulated backgrounds. The prediction is made separately for the three data-taking years, and the results are summed. The error bars on the data points represent their statistical uncertainties. The uncertainty in the predicted yields (statistical and systematic) is shown by the hatched bands.

Based on the results in this section, a systematic uncertainty of 15% is assigned in the lower two H_{T} regions, with 30% assigned in the upper H_{T} region to allow for systematic effects at the level of the larger statistical uncertainties in the closure tests in this region.

6 Systematic Uncertainties

The simulated events yields for the signal and the $\mathrm{t\bar{t}}$, Z+jets and W+jets backgrounds are affected by various systematic uncertainties. The efficiency for tagging (mis-tagging) a jet originating from two b quarks (a light flavour quark or a gluon) in simulation is corrected to match that observed in data. The uncertainty in this correction leads to an uncertainty of about 10% in the simulated signal and background event yields. The uncertainty in the soft-drop mass scale in simulation relative to data allows for a migration of events between adjacent S_i and U_i mass

regions at the level of up to 10%. The uncertainty in the simulated soft-drop mass resolution affects the widths of the simulated mass peaks. The effect is largest for signal models with small M_{H_1} and can reduce the S_i region acceptance by up to 20%. The uncertainties related to the jet energy corrections are applied to the jet properties in bins of p_{T} and η. Since there is a dedicated uncertainty for the soft-drop mass scale, their most important effect is on the event H_{T}, allowing for event migration between adjacent H_{T} bins at the level of about 4%. Integrated luminosity uncertainties of 2.5%, 2.3%, and 2.5% are applied to the 2016, 2017, and 2018 simulations, respectively [13–15]. The systematic uncertainties are assumed to be fully correlated between the data-taking years, except for the luminosity uncertainties and the 2016 double-b-tagging uncertainties, which are assumed uncorrelated. In each case, varying these correlation assumptions is found to have only a small effect on the final results. Systematic uncertainties relating to the pileup, PDFs, background cross sections, and initial-state radiation modelling were also evaluated, along with the statistical uncertainties in the simulation, and were found to give a negligible contribution to the total uncertainty.

Systematic uncertainties in the predicted multijet yields arise from the systematic uncertainties in the factors F_i. As described in Section 5, an uncertainty of 15% is applied to the F_i in the lower two H_{T} regions and 30% in the upper H_{T} region, uncorrelated between the different F_i. Except in the lowest H_{T} bin, the total uncertainties in the predicted multijet yields are dominated by the statistical uncertainties in the sideband yields $\hat{U}_i^{\mathrm{TR}}$.

7 Results

Binned maximum likelihood fits to the data in the 90 search regions (10 regions S_i per H_{T} bin for each data-taking year) are carried out under background-only and signal-plus-background hypotheses. The corresponding sideband regions U_i are fitted simultaneously, constraining the multijet contribution to the search region yields through Equation 1. The likelihoods are constructed [59] as the product of Poisson probability density functions, one for each region, with additional constraint terms for the nuisance parameters that account for the systematic uncertainties summarised in Section 6.

Figure 8 shows the yields in the search regions after the background-only fit for the combination of 2016, 2017 and 2018 data. The values and uncertainties of most nuisance parameters are unchanged by the fit, but the ones corresponding to the F_i are constrained by the data in the regions with largest yields.

Signal-plus-background fits are used to set 95% confidence level (CL) upper limits on the product of the production cross section (σ) and the $\mathrm{H}_1 \to \mathrm{b\bar{b}}$ branching ratio (BR) for the mass points considered in the benchmark signal model (Section 2). The limits are set using the modified frequentist $\mathrm{CL_s}$ criterion [60, 61], with the profile likelihood ratio test statistic [59]. The observed and expected 95% CL upper limits of $\sigma \times \mathrm{BR}$ are shown in Figure 9 as a function of M_{H_1}, for constant M_{SUSY}. The limits do not change significantly as a function of M_{SUSY} above $M_{\mathrm{SUSY}} = 2000\,\mathrm{GeV}$, which is sufficient for most signal events to populate the highest H_{T} bin (see Figure 6).

The $\sigma \times \mathrm{BR}$ upper limits are used, in conjunction with the theoretical cross section and BR values from Section 2, to exclude ranges of masses in M_{H_1} and M_{SUSY} in the benchmark model. The observed 95% CL upper limits on r, the ratio of measured and theoretical values of $\sigma \times \mathrm{BR}$, are shown in Figure 10, with the corresponding exclusion contours at $r = 1$. Masses $1200 \leq M_{\mathrm{SUSY}} \leq 2500\,\mathrm{GeV}$ are excluded within the range $40 \leq M_{\mathrm{H}_1} \leq 120\,\mathrm{GeV}$. Expected exclusion contours for the background-only scenario are also shown, and agree with the observed

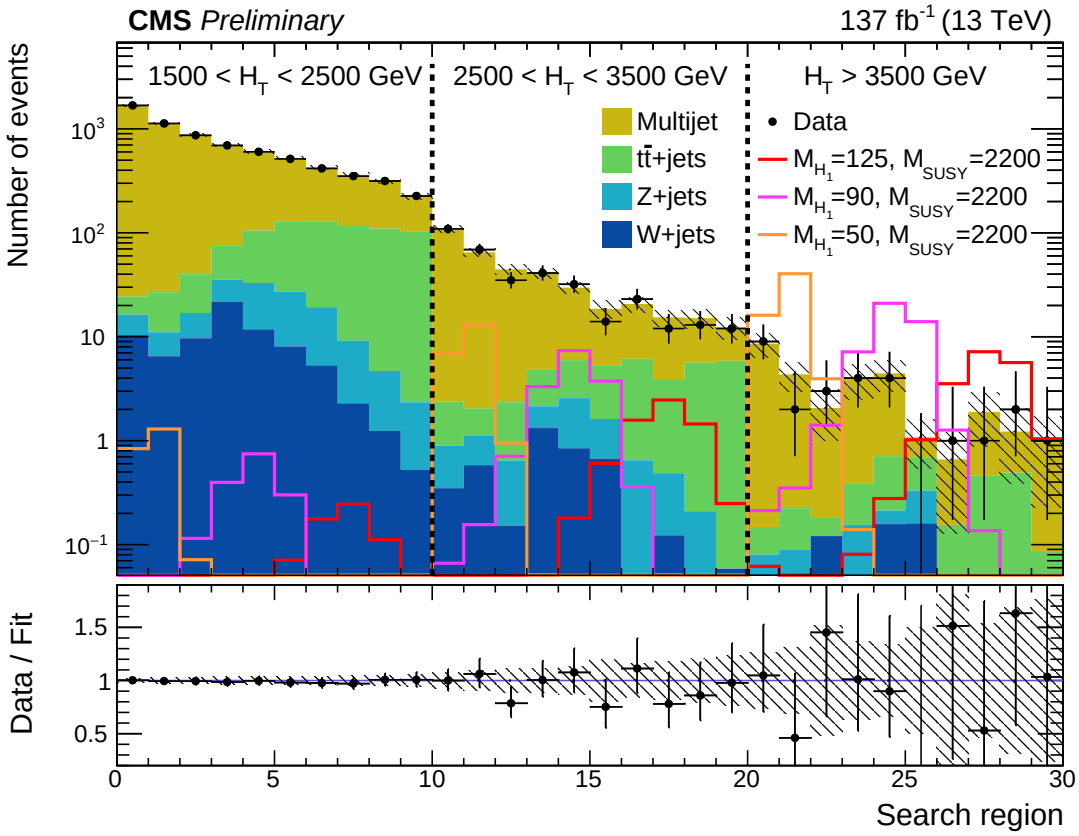

Figure 8: Search region yields after the background-only fit, summed over the three data-taking years. Example signal contributions used in the signal-plus-background fits are shown for M_{SUSY} = 2200 GeV and $M_{\mathrm{H_1}}$ = 50, 90, and 125 GeV. The numbers specified in the legend have unit GeV.

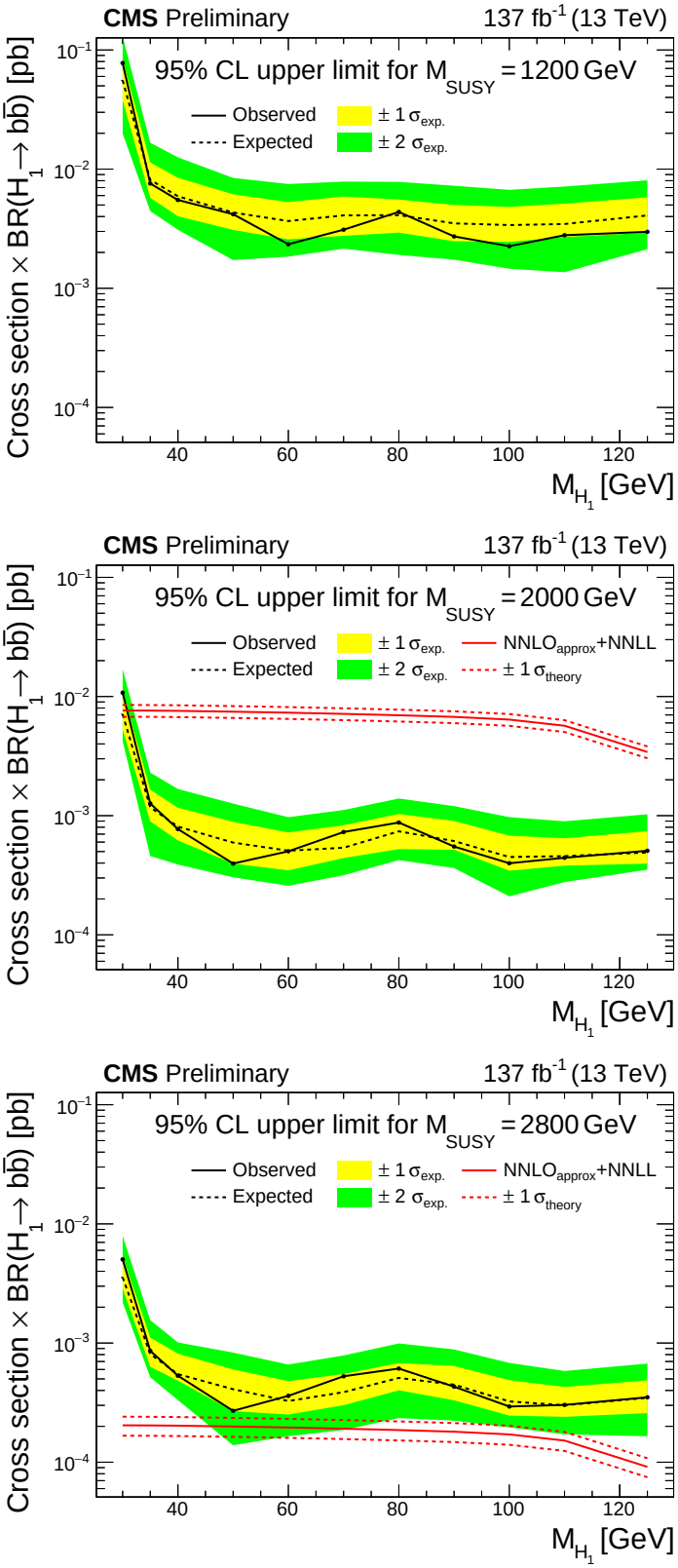

Figure 9: The observed and expected 95% CL upper limit of $\sigma \times \mathrm{BR}$ as a function of M_{H_1}. The value of M_{SUSY} is held constant at 1200 GeV (top), 2000 GeV (middle), and 2800 GeV (bottom). The solid (dashed) black line indicates the observed (median expected) limit. The yellow (green) bands indicate the expected limits with $\pm 1\sigma$ ($\pm 2\sigma$) in experimental uncertainty. The solid and dashed red lines show the theoretical prediction and its uncertainty [16–25]. The theoretical prediction is not shown in the upper plot, since its value (0.58 pb $\times$ BR) is beyond the upper extent of the vertical axis.

contours within one standard deviation. For values of M_{H_1} below 40 GeV, the distribution of signal events in the 2D soft-drop mass plane peaks outside the signal regions, resulting in a rapid reduction in signal acceptance. In the region $110 \leq M_{H_1} \leq 125$ GeV, the $H_1 \to b\bar{b}$ BR starts to decrease more quickly (see Table 2), leading to a corresponding reduction in signal sensitivity. Most of the sensitivity at large M_{SUSY} comes from the $H_T \geq 3500$ GeV region, where the statistical uncertainties in the observed yields are dominant over systematic uncertainties. This search does not explore the region outside that shown in Figure 10.

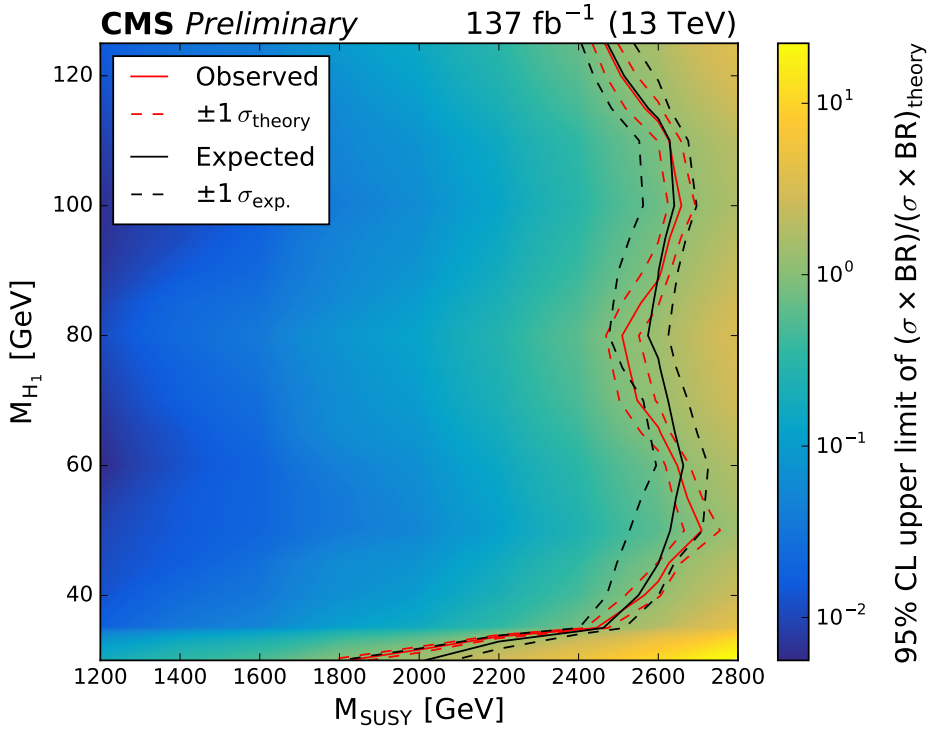

Figure 10: The observed 95% CL upper limit of $(\sigma \times \mathrm{BR})/(\sigma \times \mathrm{BR})_{\mathrm{theory}}$ (indicated by the colour scale) as a function of M_{H_1} and M_{SUSY}. The solid red (black) line delineates the observed (median expected) excluded region. The dashed red (black) line delineates the observed (expected) excluded regions with $\pm 1\sigma$ in theoretical (experimental) uncertainty.

8 Summary

This note presents a search for pairs of light Higgs bosons (H_1) produced in supersymmetric cascade decays. The targeted final states have small missing transverse momentum and two $H_1 \to b\bar{b}$ decays that are reconstructed as large-radius jets using substructure techniques. The search uses a data sample of proton-proton collision events collected by the CMS experiment at centre-of-mass energy $\sqrt{s} = 13$ TeV during LHC Run II, corresponding to an integrated luminosity of 137 fb^{-1}.

No evidence is found for any excess of events beyond the background expectations of the standard model (SM). The results are interpreted in the next-to-minimal supersymmetric extension of the SM, where a low-mass singlino leads to multi-step squark and gluino decays that can predominantly end with a boosted singlet-like H_1 boson and a low-momentum singlino-like neutralino. Upper limits are set on the product of the production cross section and the $b\bar{b}$ branching ratio of the H_1 boson for a benchmark model with almost mass-degenerate light

flavour squarks and gluinos. Under the assumption of a SM-like $H_1 \to b\bar{b}$ branching ratio, H_1 bosons with masses in the range 40–120 GeV, arising from the decays of squarks or gluinos with a mass from 1200–2500 GeV, are excluded at the 95% confidence level.

References

[1] P. Ramond, "Dual theory for free fermions", *Phys. Rev. D* **3** (1971) 2415, doi:10.1103/PhysRevD.3.2415.

[2] Y. A. Golfand and E. P. Likhtman, "Extension of the algebra of Poincaré group generators and violation of P invariance", *JETP Lett.* **13** (1971) 323.

[3] A. Neveu and J. H. Schwarz, "Factorizable dual model of pions", *Nucl. Phys. B* **31** (1971) 86, doi:10.1016/0550-3213(71)90448-2.

[4] D. V. Volkov and V. P. Akulov, "Possible universal neutrino interaction", *JETP Lett.* **16** (1972) 438.

[5] J. Wess and B. Zumino, "A Lagrangian model invariant under supergauge transformations", *Phys. Lett. B* **49** (1974) 52, doi:10.1016/0370-2693(74)90578-4.

[6] J. Wess and B. Zumino, "Supergauge transformations in four dimensions", *Nucl. Phys. B* **70** (1974) 39, doi:10.1016/0550-3213(74)90355-1.

[7] P. Fayet, "Supergauge invariant extension of the Higgs mechanism and a model for the electron and its neutrino", *Nucl. Phys. B* **90** (1975) 104, doi:10.1016/0550-3213(75)90636-7.

[8] H. P. Nilles, "Supersymmetry, supergravity and particle physics", *Phys. Rep.* **110** (1984) 1, doi:10.1016/0370-1573(84)90008-5.

[9] U. Ellwanger, C. Hugonie, and A. M. Teixeira, "The next-to-minimal supersymmetric standard model", *Physics Reports* **496** (2010) 1, doi:10.1016/j.physrep.2010.07.001.

[10] U. Ellwanger and A. Teixeira, "NMSSM with a singlino LSP: possible challenges for searches for supersymmetry at the LHC", *JHEP* **10** (2014) 113, doi:10.1007/JHEP10(2014)113.

[11] U. Ellwanger and A. Teixeira, "Excessive Higgs pair production with little MET from squarks and gluinos in the NMSSM", *JHEP* **04** (2015) 172, doi:10.1007/JHEP04(2015)172.

[12] A. Titterton et al., "Exploring sensitivity to NMSSM signatures with low missing transverse energy at the LHC", *JHEP* **10** (2018) 064, doi:10.1007/JHEP10(2018)064, arXiv:1807.10672.

[13] CMS Collaboration, "Precision luminosity measurement in proton-proton collisions at $\sqrt{s} = 13$ TeV in 2015 and 2016 at CMS", *Eur. Phys. J. C* **81** (2021) 800, doi:10.1140/epjc/s10052-021-09538-2, arXiv:2104.01927.

[14] CMS Collaboration, "CMS luminosity measurement for the 2017 data-taking period at $\sqrt{s} = 13$ TeV", CMS Physics Analysis Summary CMS-PAS-LUM-17-004, 2018.

[15] CMS Collaboration, "CMS luminosity measurement for the 2018 data-taking period at $\sqrt{s} = 13$ TeV", CMS Physics Analysis Summary CMS-PAS-LUM-18-002, 2019.

[16] W. Beenakker, R. Höpker, M. Spira, and P. M. Zerwas, "Squark and gluino production at hadron colliders", *Nucl. Phys. B* **492** (1997) 51, doi:10.1016/S0550-3213(97)00084-9, arXiv:hep-ph/9610490.

[17] A. Kulesza and L. Motyka, "Threshold resummation for squark-antisquark and gluino-pair production at the LHC", *Phys. Rev. Lett.* **102** (2009) 111802, doi:10.1103/PhysRevLett.102.111802, arXiv:0807.2405.

[18] A. Kulesza and L. Motyka, "Soft gluon resummation for the production of gluino-gluino and squark-antisquark pairs at the LHC", *Phys. Rev. D* **80** (2009) 095004, doi:10.1103/PhysRevD.80.095004, arXiv:0905.4749.

[19] W. Beenakker et al., "Soft-gluon resummation for squark and gluino hadroproduction", *JHEP* **12** (2009) 041, doi:10.1088/1126-6708/2009/12/041, arXiv:0909.4418.

[20] W. Beenakker et al., "Squark and gluino hadroproduction", *Int. J. Mod. Phys. A* **26** (2011) 2637, doi:10.1142/S0217751X11053560, arXiv:1105.1110.

[21] W. Beenakker et al., "NNLL resummation for squark-antisquark pair production at the LHC", *JHEP* **01** (2012) 076, doi:10.1007/JHEP01(2012)076, arXiv:1110.2446.

[22] W. Beenakker et al., "Towards NNLL resummation: hard matching coefficients for squark and gluino hadroproduction", *JHEP* **10** (2013) 120, doi:10.1007/JHEP10(2013)120, arXiv:1304.6354.

[23] W. Beenakker et al., "NNLL resummation for squark and gluino production at the LHC", *JHEP* **12** (2014) 023, doi:10.1007/JHEP12(2014)023, arXiv:1404.3134.

[24] W. Beenakker et al., "NNLL-fast: predictions for coloured supersymmetric particle production at the LHC with threshold and Coulomb resummation", *JHEP* **12** (2016) 133, doi:10.1007/JHEP12(2016)133, arXiv:1607.07741.

[25] A. Djouadi, J. Kalinowski, and M. Spira, "HDECAY: a program for Higgs boson decays in the standard model and its supersymmetric extension", *Computer Physics Communications* **108** (1998) 56, doi:10.1016/S0010-4655(97)00123-9, arXiv:hep-ph/9704448.

[26] J. Alwall et al., "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations", *JHEP* **07** (2014) 079, doi:10.1007/JHEP07(2014)079, arXiv:1405.0301.

[27] J. Alwall et al., "Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions", *Eur. Phys. J. C* **53** (2008) 473, doi:10.1140/epjc/s10052-007-0490-5, arXiv:0706.2569.

[28] S. Frixione, P. Nason, and G. Ridolfi, "A positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction", *JHEP* **09** (2007) 126, doi:10.1088/1126-6708/2007/09/126, arXiv:0707.3088.

[29] P. Nason, "A new method for combining NLO QCD with shower Monte Carlo algorithms", *JHEP* **11** (2004) 040, doi:10.1088/1126-6708/2004/11/040, arXiv:hep-ph/0409146.

[30] S. Frixione, P. Nason, and C. Oleari, "Matching NLO QCD computations with parton shower simulations: the POWHEG method", *JHEP* **11** (2007) 070, doi:10.1088/1126-6708/2007/11/070, arXiv:0709.2092.

[31] S. Alioli, P. Nason, C. Oleari, and E. Re, "A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX", *JHEP* **06** (2010) 043, doi:10.1007/JHEP06(2010)043, arXiv:1002.2581.

[32] NNPDF Collaboration, "Unbiased global determination of parton distributions and their uncertainties at NNLO and LO", *Nucl. Phys. B* **855** (2012) 153, doi:10.1016/j.nuclphysb.2011.09.024, arXiv:1107.2652.

[33] NNPDF Collaboration, "Parton distributions for the LHC Run II", *JHEP* **04** (2015) 040, doi:10.1007/JHEP04(2015)040, arXiv:1410.8849.

[34] NNPDF Collaboration, "Parton distributions with QED corrections", *Nucl. Phys. B* **877** (2013) 290, doi:10.1016/j.nuclphysb.2013.10.010, arXiv:1308.0598.

[35] NNPDF Collaboration, "Parton distributions from high-precision collider data", *Eur. Phys. J. C* **77** (2017) 663, doi:10.1140/epjc/s10052-017-5199-5, arXiv:1706.00428.

[36] T. Sjöstrand et al., "An introduction to PYTHIA 8.2", *Comput. Phys. Commun.* **191** (2015) 159, doi:10.1016/j.cpc.2015.01.024, arXiv:1410.3012.

[37] CMS Collaboration, "Event generator tunes obtained from underlying event and multiparton scattering measurements", *Eur. Phys. J. C* **76** (2016) 155, doi:10.1140/epjc/s10052-016-3988-x, arXiv:1512.00815.

[38] P. Skands, S. Carrazza, and J. Rojo, "Tuning PYTHIA 8.1: the Monash 2013 tune", *Eur. Phys. J. C* **74** (2014) 3024, doi:10.1140/epjc/s10052-014-3024-y, arXiv:1404.5630.

[39] CMS Collaboration, "Extraction and validation of a new set of CMS PYTHIA8 tunes from underlying-event measurements", *Eur. Phys. J. C* **80** (2020) 4, doi:10.1140/epjc/s10052-019-7499-4, arXiv:1903.12179.

[40] R. Gavin, Y. Li, F. Petriello, and S. Quackenbush, "FEWZ 2.0: A code for hadronic Z production at next-to-next-to-leading order", *Comput. Phys. Commun.* **182** (2011) 2388, doi:10.1016/j.cpc.2011.06.008, arXiv:1011.3540.

[41] R. Gavin, Y. Li, F. Petriello, and S. Quackenbush, "W physics at the LHC with FEWZ 2.1", *Comput. Phys. Commun.* **184** (2013) 208, doi:10.1016/j.cpc.2012.09.005, arXiv:1201.5896.

[42] Y. Li and F. Petriello, "Combining QCD and electroweak corrections to dilepton production in FEWZ", *Phys. Rev. D* **86** (2012) 094034, doi:10.1103/PhysRevD.86.094034, arXiv:1208.5967.

[43] M. Czakon and A. Mitov, "TOP++: a program for the calculation of the top-pair cross-section at hadron colliders", *Comput. Phys. Commun.* **185** (2014) 2930, doi:10.1016/j.cpc.2014.06.021, arXiv:1112.5675.

[44] CMS Collaboration, "Measurement of the inelastic proton-proton cross section at $\sqrt{s} = 13$ TeV", *JHEP* **07** (2018) 161, doi:10.1007/JHEP07(2018)161, arXiv:1802.02613.

[45] GEANT4 Collaboration, "GEANT4—a simulation toolkit", *Nucl. Instrum. Meth. A* **506** (2003) 250, doi:10.1016/S0168-9002(03)01368-8.

[46] CMS Collaboration, "The CMS trigger system", *JINST* **12** (2017) P01020, doi:10.1088/1748-0221/12/01/P01020, arXiv:1609.02366.

[47] CMS Collaboration, "Performance of the CMS Level-1 trigger in proton-proton collisions at $\sqrt{s} = 13$ TeV", *JINST* **15** (2020) P10017, doi:10.1088/1748-0221/15/10/P10017, arXiv:2006.10165.

[48] CMS Collaboration, "Particle-flow reconstruction and global event description with the CMS detector", *JINST* **12** (2017) P10003, doi:10.1088/1748-0221/12/10/P10003, arXiv:1706.04965.

[49] M. Cacciari, G. P. Salam, and G. Soyez, "The anti-k_T jet clustering algorithm", *JHEP* **04** (2008) 063, doi:10.1088/1126-6708/2008/04/063, arXiv:0802.1189.

[50] M. Cacciari, G. P. Salam, and G. Soyez, "FastJet user manual", *Eur. Phys. J. C* **72** (2012) 1896, doi:10.1140/epjc/s10052-012-1896-2, arXiv:1111.6097.

[51] CMS Collaboration, "Pileup mitigation at CMS in 13 TeV data", *JINST* **15** (2020) P09018, doi:10.1088/1748-0221/15/09/p09018, arXiv:2003.00503.

[52] D. Bertolini, P. Harris, M. Low, and N. Tran, "Pileup Per Particle Identification", *JHEP* **10** (2014) 059, doi:10.1007/JHEP10(2014)059, arXiv:1407.6013.

[53] CMS Collaboration, "Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV", *JINST* **12** (2017) P02014, doi:10.1088/1748-0221/12/02/P02014, arXiv:1607.03663.

[54] CMS Collaboration, "Jet energy scale and resolution performance with 13 TeV data collected by CMS in 2016", Detector Performance Report CMS-DP-2018-028, 2018.

[55] CMS Collaboration, "Identification of double-b quark jets in boosted event topologies", CMS Physics Analysis Summary CMS-PAS-BTV-15-002, 2016.

[56] T. Adams et al., "Beam test evaluation of electromagnetic calorimeter modules made from proton-damaged $PbWO_4$ crystals", *JINST* **11** (2016) P04012, doi:10.1088/1748-0221/11/04/P04012.

[57] A. J. Larkoski, S. Marzani, G. Soyez, and J. Thaler, "Soft Drop", *JHEP* **05** (2014) 146, doi:10.1007/JHEP05(2014)146, arXiv:1402.2657.

[58] M. Dasgupta, A. Fregoso, S. Marzani, and G. P. Salam, "Towards an understanding of jet substructure", *JHEP* **09** (2013) 029, doi:10.1007/JHEP09(2013)029, arXiv:1307.0007.

[59] ATLAS and CMS Collaborations, "Procedure for the LHC Higgs boson search combination in summer 2011", ATLAS/CMS joint note ATL-PHYS-PUB-2011-011, CMS-NOTE-2011-005, 2011.

[60] T. Junk, "Confidence level computation for combining searches with small statistics", *Nucl. Instrum. Meth. A* **434** (1999) 435, doi:10.1016/S0168-9002(99)00498-2, arXiv:hep-ex/9902006.

[61] A. L. Read, "Presentation of search results: The CL_s technique", *J. Phys. G* **28** (2002) 2693, doi:10.1088/0954-3899/28/10/313.